

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 20:46:52 ; Search time 8673 seconds

(without alignments)
17519.399 Million cell updates/sec

Title: US-09-918-187-3

Perfect score: 5221
Sequence: 1 ataaagggggctgagga.....aatctaaaaaaaaaaaaa 5221

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 993650

Minimum DB seq length: 8
Maximum DB seq length: 100

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

GenB1:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-----------|--------------------|
| 1 | 39 | 0.7 | 39 | 6 | AR003340 | AR003340 Sequence |
| 2 | 33.6 | 0.6 | 86 | 3 | AY009729 | AY009729 Boophilus |
| 3 | 32.4 | 0.6 | 90 | 4 | OA05003 | AJ005003 Ovis ari |
| 4 | 31.6 | 0.6 | 96 | 9 | H0MDINOCJ | L21846 Human dlnuc |
| 5 | 31.2 | 0.6 | 90 | 10 | RNDI12243 | U12243 Rattus norv |
| 6 | 31 | 0.6 | 69 | 4 | AF339942 | AF339942 Puma conc |
| 7 | 31 | 0.6 | 69 | 4 | AF339942 | AF339942 Puma conc |
| 8 | 30.6 | 0.6 | 83 | 5 | AF420492 | AF420492 Salmo sal |
| 9 | 30.4 | 0.6 | 71 | 4 | AF055530 | AF055530 Didelphis |
| 10 | 30.4 | 0.6 | 73 | 4 | AF339937 | AF339937 Puma conc |
| 11 | 30.2 | 0.6 | 88 | 3 | AY009730 | AY009730 Boophilus |
| 12 | 30 | 0.6 | 63 | 6 | AX340088 | AX340088 Sequence |
| 13 | 30 | 0.6 | 73 | 11 | AF042300 | AF042300 Homo sapi |
| 14 | 29.6 | 0.6 | 92 | 11 | G68120 | G68120 DRS3168 (CR |
| 15 | 29.6 | 0.6 | 93 | 6 | I05524 | I05524 Sequence 11 |
| 16 | 29.6 | 0.6 | 93 | 6 | AF126190 | AF126190 Cervus da |
| 17 | 29.4 | 0.6 | 73 | 6 | AX397053 | AX397053 Sequence |
| 18 | 29.4 | 0.6 | 95 | 11 | AU026788 | AU026788 Rattus no |
| 19 | 29.2 | 0.6 | 97 | 8 | AB055133 | AB055133 Polynomum |
| 20 | 29 | 0.6 | 97 | 4 | AF213233 | AF213233 Bison bis |
| 21 | 29 | 0.6 | 99 | 11 | AU049997 | AU049997 Rattus no |
| 22 | 28.8 | 0.6 | 50 | 6 | I31455 | I31455 Sequence 36 |
| 23 | 28.8 | 0.6 | 65 | 6 | I31204 | I31204 Sequence 11 |
| 24 | 28.8 | 0.6 | 65 | 6 | I31509 | I31509 Sequence 42 |
| 25 | 28.8 | 0.6 | 75 | 11 | AF042301 | AF042301 Homo sapi |
| 26 | 28.8 | 0.6 | 77 | 11 | AF042302 | AF042302 Homo sapi |
| 27 | 28.8 | 0.6 | 91 | 3 | AY009760 | AY009760 Boophilus |
| 28 | 28.6 | 0.5 | 87 | 9 | HUML | L23535 Human Chrom |
| 29 | 28.6 | 0.5 | 87 | 11 | G31698 | G31698 SWS1885 Er |
| 30 | 28.6 | 0.5 | 89 | 11 | AU027741 | AU027741 Rattus no |
| 31 | 28.6 | 0.5 | 97 | 4 | MAA309850 | AJ309850 Martes ma |
| 32 | 28.4 | 0.5 | 73 | 11 | AF042303 | AF042303 Homo sapi |
| 33 | 28.4 | 0.5 | 91 | 10 | MMNDMS20 | X55238 M.musculus |
| 34 | 28.4 | 0.5 | 93 | 4 | AF213232 | AF213232 Bison bis |
| 35 | 28.2 | 0.5 | 51 | 6 | I31425 | I31425 Sequence 33 |
| 36 | 28.2 | 0.5 | 60 | 3 | DROKS2 | M35655 D.melanogas |
| 37 | 28.2 | 0.5 | 62 | 4 | AF1851152 | AF185112 Lasiorhin |
| 38 | 28.2 | 0.5 | 69 | 11 | AF042298 | AF042298 Homo sapi |
| 39 | 28.2 | 0.5 | 85 | 10 | MMNDMS16 | X55234 M.musculus |
| 40 | 28.2 | 0.5 | 86 | 6 | AX203294 | AX203294 Sequence |
| 41 | 28.2 | 0.5 | 92 | 5 | AF019993 | AF019993 Clupea pa |
| 42 | 28.2 | 0.5 | 93 | 4 | AF213195 | AF213195 Bison bis |
| 43 | 28 | 0.5 | 79 | 3 | AY009762 | AY009762 Boophilus |
| 44 | 28 | 0.5 | 83 | 11 | AF042305 | AF042305 Homo sapi |
| 45 | 28 | 0.5 | 87 | 11 | HUWUT792A | L31205 Human STR u |

ALIGNMENTS

RESULT 1
LOCUS AR003340 39 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 62 from patent US 5744300.
ACCESSION AR003340
VERSION AR003340.1 GI:3964599
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 39)
AUTHORS Linskens,M.H.K., Hirsch,K.S., Villeponteau,B., Feng,J., Funk,W. and West,M.David.
TITLE Methods and reagents for the identification and regulation of senescence-related genes

Pred. No. is the number of results predicted by chance to have a

[illegible][illegible]

| LOCUS | NCBI 2243/c | 90 bp | DNA | linear | ROD 30-JAN-1996 |
|------------|---|-----------|-----|--------|-----------------|
| DEFINITION | Rattus norvegicus LEW/N clone D10AR305A satellite DNA sequence. | | | | |
| ACCESSION | U12243 | | | | |
| VERSION | U12243.1 | GI:529178 | | | |
| KEYWORDS | | | | | |
| SOURCE | | | | | |
| ORGANISM | Rattus norvegicus. | | | | |
| REFERENCE | Rattus norvegicus. | | | | |
| AUTHORS | Eukariyoti; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| TITLE | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; | | | | |
| REFERENCE | Rattus. | | | | |
| AUTHORS | 1 (bases 1 to 90) | | | | |
| TITLE | Du,Y., Remmers,E.F., Zha,H., Mathern,P. and Wilder,R.L. | | | | |
| REFERENCE | Simple sequence repeats identified in LEW/N rat genomic library | | | | |
| AUTHORS | unpublished | | | | |
| TITLE | 2 (bases 1 to 90) | | | | |
| REFERENCE | Zha,H., Remmers,E.F., Du,Y., Cash,J.M., Goldmuntz,E.A., | | | | |
| AUTHORS | Crofford,L.J. and Wilder,R.L. | | | | |
| TITLE | The rat athletic nude (nu) locus is closely linked to the inducible | | | | |
| REFERENCE | nitric oxide synthase gene (NOS2) | | | | |
| AUTHORS | Mamm. Genome 6 (2), 137-138 (1995) | | | | |
| TITLE | 3 (bases 1 to 90) | | | | |
| REFERENCE | Du,Y. | | | | |
| AUTHORS | Direct Submission | | | | |
| TITLE | Submitted (12-JUL-1994) Ying Du, National Institutes of Health, | | | | |
| REFERENCE | ARB/NIAMS, 9000 Rockville Pike, Bethesda, MD 20892, USA | | | | |
| AUTHORS | Location/Qualifiers | | | | |
| TITLE | 1..90 | | | | |
| REFERENCE | /organism="Rattus norvegicus" | | | | |
| AUTHORS | /strain="LEW/N" | | | | |
| TITLE | /db_xref="taxon:10116" | | | | |
| REFERENCE | /clone="D10AR305A" | | | | |
| AUTHORS | /cell_type="hepatocyte" | | | | |
| TITLE | /tissue_type="liver" | | | | |
| REFERENCE | /clone_lib="Mbol restricted genomic DNA" | | | | |
| AUTHORS | 29..81 | | | | |
| TITLE | BASE COUNT 35 a 33 c 6 g 15 t 1 others | | | | |
| REFERENCE | ORIGIN | | | | |
| AUTHORS | Query Match 0.6% Score 31.2; DB 10; Length 90; | | | | |
| TITLE | - Best Local Similarity 70.0% Pred. No. 1.1e+04; | | | | |
| REFERENCE | Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0; | | | | |
| AUTHORS | QY 4078 TGTATGTGTGCTGTGGGTGGGTGGGAGTGTCTCTGATGAAGACAGATTTC 4137 | | | | |
| TITLE | | | | | |
| REFERENCE | DB 67 TGTAC 8 | | | | |
| AUTHORS | RESULT 6 | | | | |
| TITLE | AF339940/c | | | | |
| REFERENCE | LOCUS AF339940 69 bp DNA linear MAM 12-JUN-2001 | | | | |
| AUTHORS | DEFINITION Puma concolor clone Pco553 locus FCA090 microsatellite sequence. | | | | |
| TITLE | ACCESSION AF339940 | | | | |
| REFERENCE | VERSION AF339940.1 GI:14348695 | | | | |
| AUTHORS | KEYWORDS | | | | |
| TITLE | SOURCE | | | | |
| REFERENCE | ORGANISM | | | | |
| AUTHORS | Puma concolor. | | | | |
| TITLE | Puma concolor | | | | |
| REFERENCE | Eukariyoti; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| AUTHORS | Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Puma. | | | | |
| TITLE | 1 (bases 1 to 69) | | | | |
| REFERENCE | Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J. | | | | |
| AUTHORS | Patterns of size homoplasy at 10 microsatellite loci in pumas (Puma | | | | |
| TITLE | concolor) | | | | |
| REFERENCE | MOJ. Biol. Evol. 18 (6), 1151-1156 (2001) | | | | |
| AUTHORS | JOURNAL | | | | |
| TITLE | MEDLINE | | | | |
| REFERENCE | PUBMED | | | | |
| AUTHORS | 2 (bases 1 to 69) | | | | |
| TITLE | Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J. | | | | |
| REFERENCE | Direct Submission | | | | |
| AUTHORS | Submitted (24-JAN-2001) Fisheries and Wildlife Sciences, VPI&SU, | | | | |

| | |
|-----------------------|---|
| FEATURES | 134 Cheatham, Blacksburg, VA 24061-0321, USA |
| SOURCE | Location/Qualifiers 1..69 /organism="Puma concolor" /db_xref="taxon:9696" /chromosome="A1" /map="Locus FCA090" /clone="PCo553" |
| repeat_region | 1..69 /note="microsatellite" /rpl_type=tandem |
| BASE COUNT | 33 a 24 c 6 g 6 t |
| ORIGIN | |
| Query Match | 0.6% ; Score 31; DB 4; Length 69; |
| Best Local Similarity | 72.7%; Pred. No. 1.2e+04; |
| Matches | 40; Conservative 0; Mismatches 15; Indels 0; Gaps 0; |
| Dn | 57 TCAGGCTGAGGGCCCCCATGTATGTGGTGGGTGGGTGGGTGGGTGTCT 4114 |
| RESULT 7 | |
| LOCUS | AF339942 69 bp DNA linear MAR 12-JUN-2001 |
| DEFINITION | Puma concolor clone PCo564 locus FCA090 microsatellite sequence. |
| ACCESSION | AF339942 |
| VERSION | AF339942.1 GI:14348697 |
| KEYWORDS | ' Puma concolor. |
| SOURCE | Puma concolor |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Puma. Culver,M.; Menotti-Raymond,M.A. and O'Brien,S.J. Patterns of size homoplasy at 10 microsatellite loci in pumas (Puma concolor) Mol. Biol. Evol. 18 (6), 1151-1156 (2001) |
| JOURNAL | MOL. BIOL. EVOL. 18 (6), 1151-1156 (2001) |
| MEDLINE | 21265424 |
| PUBMED | 11371606 |
| REFERENCE | 2 (bases 1 to 69) Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J. Direct Submission Submitted (24-JAN-2001) Fisheries and Wildlife Sciences, VPI&SU, 134 Cheatham, Blacksburg, VA 24061-0321, USA |
| AUTHORS | |
| TITLE | |
| JOURNAL | |
| FEATURES | Location/Qualifiers 1..69 /organism="Puma concolor" /db_xref="taxon:9696" /chromosome="A1" /map="Locus FCA090" /clone="PCo564" |
| SOURCE | 1..69 /note="microsatellite" /rpl_type=tandem |
| BASE COUNT | 33 a 24 c 6 g 6 t |
| ORIGIN | |
| Query Match | 0.6% ; Score 31; DB 4; Length 69; |
| Best Local Similarity | 72.7%; Pred. No. 1.2e+04; |
| Matches | 40; Conservative 0; Mismatches 15; Indels 0; Gaps 0; |
| OY | 4060 TCAGGCTGAGGGCCCCCATGTATGTGGTGGGTGGGTGGGTGGGTGTCT 4114 |
| Dn | 57 TCATTACTTAGAGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCT 3 |
| RESULT 8 | |
| LOCUS | AF420492 83 bp DNA linear VRT 03-OCT-2001 |
| DEFINITION | Salmo salar clone SaalNUG microsatellite sequence. |
| ACCESSION | AF420492 |

| | | | | |
|-----------------------|---|--|----------------|-----------|
| oy | 4753 | TA9A | 4756 | |
| Db | 65 | CAAA | 68 | |
| RESULT 10 | AF339937/c | 73 bp | DNA | linear |
| LOCUS | Puma concolor clone Pco036 | locus FCA090 | microsatellite | sequence. |
| DEFINITION | AF339937 | | | |
| ACCESSION | AF339937.1 | GI:14348692 | | |
| VERSION | | | | |
| KEYWORDS | | | | |
| SOURCE | Puma concolor. | | | |
| ORGANISM | Puma concolor | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| AUTHORS | Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Puma. | | | |
| TITLE | 1 (bases 1 to 73) | | | |
| JOURNAL | Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J. | | | |
| MEDLINE | Patterns of size homoplasy at 10 microsatellite loci in pumas (Puma concolor) | | | |
| PUBMED | Mol. Biol. Evol. 18 (6), 1151-1156 (2001) | | | |
| REFERENCE | 21265424 | | | |
| AUTHORS | 11371606 | | | |
| TITLE | 2 (bases 1 to 73) | | | |
| JOURNAL | Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J. | | | |
| FEATURES | Direct Submission | | | |
| source | Submitted (24-JUN-2001) Fisheries and Wildlife Sciences, VPI&SU, 134 Cheatam, Blacksburg, VA 24061-0321, USA | | | |
| repeat_region | location/Qualifiers | | | |
| BASE COUNT | 1..73 | | | |
| ORIGIN | /organism="Puma concolor" | | | |
| | /db_xref="taxon:9696" | | | |
| | /chromosome="A1" | | | |
| | /map="locus FCA090" | | | |
| | /clone="Pco036" | | | |
| | 1..73 | | | |
| | /note="microsatellite" | | | |
| | /rpt_type=tandem | | | |
| | 35 a 26 c 6 g | | | |
| Query Match | 0.6%; Score 30.4; DB 4; Length 73; | | | |
| Best local Similarity | 71.4%; Pred. No. 1,8e+04; | | | |
| Matches | 40; Conservative 0; Mismatches 16; Indels 0; Gaps 0; | | | |
| oy | 4060 | TCAGGCTGAGGCCCCCATGATGATGTGTCGTCGGTGGTGGGAGTGTCTG | 4115 | |
| Db | 61 | TCATACTTAGAGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG | 6 | |
| RESULT 11 | AY009730 | 88 bp | DNA | linear |
| LOCUS | Boophilus microplus clone PNC 75 | isolate Eacham3R | microsatellite | sequence. |
| DEFINITION | AY009730 | | | |
| ACCESSION | AY009730.1 | GI:13603306 | | |
| VERSION | | | | |
| KEYWORDS | | | | |
| SOURCE | Boophilus microplus. | | | |
| ORGANISM | Boophilus microplus | | | |
| REFERENCE | Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari; | | | |
| AUTHORS | Parasitiformes; Ixodida; Ixodidae; Boophilus. | | | |
| TITLE | 1 (bases 1 to 88) | | | |
| JOURNAL | Chitgare,N.N. and Barker,S.C. | | | |
| REFERENCE | Microsatellite DNA population genetic analysis of Boophilus microplus (Acari:Ixodidae) | | | |
| AUTHORS | Unpublished | | | |
| TITLE | 2 (bases 1 to 88) | | | |
| JOURNAL | Chitgare,N.N. and Barker,S.C. | | | |
| REFERENCE | Direct Submission | | | |
| AUTHORS | Submitted (11-OCT-2000) Parasitology and Microbiology, The University of Queensland, St. Lucia Campus, Brisbane, Queensland | | | |

```

FEATURES          QLD 4072, Australia
SOURCE            1.88      Location/Qualifiers
                  /organism="Boophilus microplus"
                  /isolate="Bacham38"
                  /db_xref="taxon:6941"
                  /clone="PNC 75"
                  /country="Australia"
                  repeat_region 1.88
                  /note="microsatellite"
                  /rpl_type="tandem
                  /rpl_unit="ca

BASE COUNT       9 a      8 c      35 g      36 t
ORIGIN

Query Match      0.6%; Score 30.2; DB 3; Length 88;
Best Local Similarity 62.7%; Pred. No. 2e+04;
Matches 47; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 4038 CTACATAGAGCTGATGACCTGAGGCTGAGGCCCAATGATGCTGGCTGGGCTG 4097
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 5 CGACATGAGAGTGTGCTACTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 64

QY 4098 TGGGTGGGAGCTGTGT 4112
      || || | | | | | | | |
DB 65 TGTGTGTGTGTGTGTGT 79

RESULT 12
AX340088          AX340088      63 bp      DNA      linear      PAT 10-JAN-2002
LOCUS             Sequence 335 from Patent WO016388.
DEFINITION        AX340088
ACCESSION         AX340088.1 GI:18136069
VERSION           human.
KEYWORDS          Homo sapiens
SOURCE            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
ORGANISM          human.

REFERENCE
AUTHORS           Jiang, Y., Harlocker, S.L. and Secrist, H.
TITLE             Compositions and methods for the therapy and diagnosis of colon
                  cancer
JOURNAL           Patent: WO 0196388-A 335 20-DEC-2001;
FEATURES          CORIXA CORPORATION (US)
SOURCE            Location/Qualifiers
                  1.63
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"

BASE COUNT       9 a      2 c      4 g      44 t      4 others
ORIGIN

Query Match      0.6%; Score 30; DB 6; Length 63;
Best Local Similarity 72.5%; Pred. No. 2.2e+04;
Matches 37; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

QY 4475 TAAATTGAATGATTTGATTAATAGCTTTTTTTTTTTGGGCTTAAAG 4525
      || || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 11 TTTNTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTGGTTAAAD 61

RESULT 13
AF042300/c        AF042300      73 bp      DNA      linear      STS 26-JAN-1999
LOCUS             Homo sapiens chromosome 6 map 6p21.3, sequence tagged site.
DEFINITION        AF042300
ACCESSION         AF042300.1 GI:4191287
VERSION           STS.
KEYWORDS          Homo sapiens
SOURCE            Homo sapiens
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 73)

```

```

AUTHORS           Lin, L., Jin, L., Kimura, A., Carlington, M. and Mignot, E.
TITLE             DQ microsatellite association studies in three ethnic groups
JOURNAL           Tissue Antigens 50 (5), 507-520 (1997)
MEDLINE           98049194
PUBMED           9389326
REFERENCE         2 (bases 1 to 73)
AUTHORS           Lin, L., Jin, L., Lin, X., Voros, A., Underhill, P. and Mignot, E.
TITLE             Microsatellite single nucleotide polymorphisms in the HLA-DQ region
JOURNAL           Tissue Antigens 52 (1), 9-18 (1998)
MEDLINE           98378271
PUBMED           9714469
REFERENCE         3 (bases 1 to 73)
AUTHORS           Lin, L., Jin, L., Lin, X., Voros, A., Underhill, P. and Mignot, E.
TITLE             Direct Submission
JOURNAL           Submitted (12-JAN-1998) Psychiatry and Behavioral Sciences,
                  Stanford University, Lab Surge Building, 1201 Welch Road, Room
                  1126, Palo Alto, CA 94304, USA
FEATURES          Location/Qualifiers
SOURCE            1.73
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="6"
                  /map="6p21.3"
                  repeat_region 1.73
                  /note="microsatellite"
                  /rpl_type="tandem
                  /rpl_unit="ca

BASE COUNT       32 a      29 c      3 g      9 t
ORIGIN

Query Match      0.6%; Score 30; DB 11; Length 73;
Best Local Similarity 67.7%; Pred. No. 2.3e+04;
Matches 42; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 4064 GCTGAGGAGCCCAATGATGCTGGCTGGGCTGGGAGTGTCTGCTGACTAA 4123
      || || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 64 GTTGTGCAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTAATATA 5

QY 4124 GG 4125
      ||
DB 4 GG 3

RESULT 14
G68120/c          G68120      92 bp      DNA      linear      STS 23-APR-2001
LOCUS             D7S3168 (CKG2) p1 library (Shepherd et al., 1994) Homo sapiens STS
DEFINITION        genomic clone p1317a1, sequence tagged site.
ACCESSION         G68120
VERSION           G68120.1 GI:13775132
KEYWORDS          STS.
SOURCE            Homo sapiens
ORGANISM          Homo sapiens
SOURCE            Homo sapiens
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE         1 (bases 1 to 92)
AUTHORS           Heus, H.C., Hing, A., van Baren, M.J., Joosse, M., Breedveld, G.J.,
                  Wang, J.C., Burgess, A., Donnis-Keller, H., Berglund, C., Zgurics, J.,
                  Scherer, S.W., Rommens, J.M., Oostra, B.A. and Heutink, P.
TITLE             A physical and transcriptional map of the preaxial polydactyly
                  locus on chromosome 7q36
JOURNAL           Genomics 57 (3), 342-351 (1999)
MEDLINE           99263496
PUBMED           10329000
COMMENT           GDB: 10797894
                  GDB_DSEG: D7S3168
                  Contact: Scherer, S.W.; Rommens, J.M.; Tsui, L.C.; Heus, H.C.;
                  Oostra, B.A.; Heutink, P.
                  The Human Chromosome 7 Project
                  Department of Genetics, Hospital for Sick Children
                  555 University Avenue, Toronto, Ontario M5G 1X8, Canada
                  Tel: 416 813 7613

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: November 23, 2002, 20:46:07 : Search time 659 Seconds
(without alignments)
17841.704 Million cell updates/sec

Title: US-09-918-187-3

Perfect score: 5221
Sequence: 1 ataaagggggtctgagaa.....atctaaaaaataaaaaa 5221

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2389434

Minimum DB seq length: 8
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001C.DAT.*
- 24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the distribution, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 52.4 | 1.0 | 54 | 17 | AAT28167 |
| 2 | 49.4 | 0.9 | 51 | 22 | AAL28163 |
| 3 | 39 | 0.7 | 39 | 17 | AAT28105 |
| 4 | 36.6 | 0.7 | 67 | 21 | AAAB7350 |
| 5 | 36.2 | 0.7 | 65 | 24 | ABNS3617 |
| 6 | 34.2 | 0.7 | 97 | 16 | AAT21227 |
| 7 | 33.8 | 0.6 | 83 | 24 | ABNS2844 |
| 8 | 31 | 0.6 | 95 | 17 | AAT13984 |
| 9 | 30.6 | 0.6 | 93 | 24 | ABNS7931 |

| | | | | | | | |
|---|----|------|-----|-----|----|-----------|---------------------|
| C | 10 | 30.4 | 0.6 | 64 | 22 | AAH48695 | Human G-protein su |
| | 11 | 30 | 0.6 | 63 | 24 | ABL36746 | Human colon tumour |
| | 12 | 30 | 0.6 | 65 | 24 | ABNS28832 | Rat spliced trans |
| | 13 | 29.8 | 0.6 | 68 | 21 | AACT17498 | Human secreted pro |
| | 14 | 29.6 | 0.6 | 60 | 24 | ABNS36661 | Human spliced tran |
| | 15 | 29.4 | 0.6 | 73 | 24 | ABNS45717 | cdNA encoding colo |
| | 16 | 29.2 | 0.6 | 90 | 12 | AAQ11762 | Self-complementary |
| | 17 | 29 | 0.6 | 95 | 24 | ABNS19692 | Human genome-deriv |
| | 18 | 28.8 | 0.6 | 50 | 18 | AAT65787 | Repeat sequence fr |
| | 19 | 28.8 | 0.6 | 65 | 18 | AAT65724 | Repeat sequence fr |
| | 20 | 28.8 | 0.6 | 65 | 18 | AAT65072 | (dc-d)n.(dc-d)n |
| | 21 | 28.8 | 0.6 | 75 | 21 | AAT29909 | Human secreted pro |
| | 22 | 28.2 | 0.5 | 51 | 18 | AAT65777 | Repeat sequence fr |
| | 23 | 28.2 | 0.5 | 86 | 22 | AA523317 | Human prostate can |
| | 24 | 28 | 0.5 | 62 | 16 | AAT20450 | Human gene signatu |
| | 25 | 27.8 | 0.5 | 57 | 18 | AAT65792 | Repeat sequence fr |
| | 26 | 27.6 | 0.5 | 73 | 13 | AAQ34140 | Sequence of a micr |
| | 27 | 27.6 | 0.5 | 74 | 24 | AAI69128 | Activated T-cell d |
| | 28 | 27.6 | 0.5 | 87 | 21 | AACT18326 | Human secreted pro |
| | 29 | 27.4 | 0.5 | 75 | 19 | AAV27232 | Human gene sequence |
| | 30 | 27.4 | 0.5 | 75 | 22 | AAH43150 | seqIPD#608P51. S |
| | 31 | 27.2 | 0.5 | 50 | 18 | AAT65774 | Repeat sequence fr |
| | 32 | 27.2 | 0.5 | 55 | 24 | ABNS3132 | HIV detection meth |
| | 33 | 27.2 | 0.5 | 55 | 24 | AAI45494 | HIV-1 RNA capture |
| | 34 | 27.2 | 0.5 | 50 | 22 | AA523666 | Tetracycline promo |
| | 35 | 27 | 0.5 | 85 | 19 | AAV32418 | Homo sapiens clone |
| | 36 | 27 | 0.5 | 85 | 22 | AAV98434 | Human CDNA clone B |
| | 37 | 27 | 0.5 | 87 | 24 | AAI69175 | Activated T-cell d |
| | 38 | 27 | 0.5 | 99 | 21 | AACT11621 | Human secreted pro |
| | 39 | 27 | 0.5 | 100 | 15 | AAO62573 | Candida albicans-s |
| | 40 | 26.8 | 0.5 | 46 | 18 | AAT65709 | Repeat sequence fr |
| | 41 | 26.8 | 0.5 | 49 | 18 | AAT65785 | Repeat sequence fr |
| | 42 | 26.8 | 0.5 | 53 | 18 | AAT65768 | Repeat sequence fr |
| | 43 | 26.8 | 0.5 | 54 | 13 | AAQ33870 | Microsatellite seq |
| | 44 | 26.8 | 0.5 | 56 | 18 | AAT65063 | (dc-d)n.(dc-d)n |
| | 45 | 26.8 | 0.5 | 65 | 18 | AAT65730 | Repeat sequence fr |

ALIGNMENTS

| | |
|------------|---|
| RESULT 1 | |
| AAT28167/c | |
| ID | AAT28167 standard; DNA; 54 BP. |
| XX | |
| AC | AAT28167; |
| XX | |
| DT | 06-JAN-1997 (first entry) |
| XX | |
| DE | Senescence-related gene sequence 10P1. |
| XX | |
| KW | Human: fibroblast; AIDS: enhanced differential display; mRNA preparation; |
| KW | senescent cell; quiescent cell; dividing cell; senescence-related gene; |
| KW | gene expression; non-senescent cell; age-related lipofuscin; retina; EDO; |
| KW | therapy; liver spot; donor tissue; senescent melanocyte; melanin; |
| KW | hypopigmentation; ss. |
| OS | |
| XX | Synthetic. |
| PN | |
| XX | W09613610-A2. |
| PD | |
| XX | 09-MAY-1996. |
| PF | |
| XX | 24-AUG-1995; 95WO-US11230. |
| PR | |
| XX | 31-OCT-1994; 94US-0332420. |
| PA | |
| XX | (GERO-) GERON CORP. |
| PI | |
| XX | Feng J, Funk W, Hirsch KS, Linskens MK, Villeponteau B; |
| XX | West MD; |
| DR | WPI; 1996-251464/25. |

```

XX Identifying, isolating and regulating senescence-related genes -
PT useful to ameliorate problems associated with accumulation of
PR senescent cells, e.g. age-related lipofuscin accumulation in the
PT retina and AIDS
PS
PS Claim 8; Page 51; 135pp; English.
XX
CC ART28076-128113, and ART28131-T28173 represent novel senescent-related
CC gene sequences isolated from fibroblasts using the method of the
CC invention. In the method of the invention, mRNA is isolated from a
CC senescent cell, and a young quiescent cell, and the mRNAs are amplified
CC (using primers such as those shown in ART28044-T28075) in separate
CC reaction mixtures. The amplified sequences are then separated by size or
CC charge, and the products are analysed to identify a gene from young
CC quiescent cells and dividing cells, that is present at a different level
CC from senescent cells. To enhance the method even more, it can be
CC performed in conjunction with an enhanced fibroblastal display (EDD)
CC method (an mRNA preparation method) on the fibroblasts. The method can
CC be used for the rapid and efficient identification and isolation of
CC senescence-related genes and gene products, and to detect and distinguish
CC between senescent and non-senescent cells. It can also be used to
CC destroy cells expressing senescence specific (or related) gene products,
CC and to screen for compounds capable of altering gene expression in
CC senescent cells. The method can also be used to ameliorate problems
CC associated with the accumulation of senescent cells such as age-related
CC lipofuscin accumulation in the retina, and in the treatment of AIDS.
CC Also, the method can be used to distinguish young cells from senescent
CC cells in donor tissue, which is useful in removing senescent melanocytes
CC overexpressing melanin which cause hypopigmentation, or liver spots.
XX
XX Sequence 54 BP; 17 A; 9 C; 9 G; 19 T; 0 other;
SO
Query Match 1.0%; Score 52.4; DB 17; Length 54;
Best Local Similarity 98.1%; Pred. No. 0.00039;
Matches 53; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1402 AGGATGCCTAAAGATGATGTATTGAACCACTTCAGTACAGATTTCTTTAAAT 1455
DB 54 AGGATGCCAAGAAGATGATCATGTTAACCACTTCAGTACAGATTTCTTTAAAT 1
RESULT 2
AL28163
AL28163 standard; DNA; 51 BP.
XX
XX AL28163;
DT 24-JAN-2002 (first entry)
XX
XX Human SNP oligonucleotide #1371.
DE
RW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
RW neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;
RW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
RW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
RW complement related protein; cytochrome; kinesin; cytokine; interferon;
RW interleukin; G-protein coupled receptor; thioesterase; inflammation;
RW multifactorial disease; autoimmune disease; infection;
RW nervous system disease; ss.
XX
OS Homo sapiens.
PN WO200147944-A2.
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
PA (CURA-) CURAGEN CORP.
```

| | |
|--------------------------|---|
| XX | Shinkets RA, Leach M; |
| PI | WPI; 2001-465210/50. |
| DR | |
| XX | |
| XX | Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases, |
| PT | oncogenes and histones, useful for diagnosing and treating, e.g. |
| PI | cancer, autoimmune diseases and infections - |
| PS | Claim 1; Page 1772; 4143pp; English. |
| XX | |
| CC | The present invention relates to oligonucleotides encoding polymorphic |
| CC | variants of proteins related to amylases, amyloid proteins, angiotensin, |
| CC | apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes, |
| CC | histones, kinases, colony stimulating factors, complement related |
| CC | proteins, cytochromes, kinesins, cytokines, interferons, interleukins, |
| CC | G-protein coupled receptors and thioesterases. The present sequence is |
| CC | one such oligonucleotide. The oligonucleotides and the peptides encoded |
| CC | by them may be used in the prevention, diagnosis and treatment of |
| CC | diseases associated with inappropriate expression of the proteins listed |
| CC | above. Disorders that may be prevented, diagnosed and/or treated include |
| CC | multifactorial diseases with a genetic component, such as autoimmune |
| CC | diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes, |
| CC | systemic lupus erythematosus and Grave's disease), inflammation, cancer |
| CC | (e.g. cancers of the bladder, brain, breast, colon and kidney, |
| CC | leukemia), diseases of the nervous system and an infection of pathogenic |
| CC | organisms. |
| CC | |
| SO | Sequence 51 BP; 11 A; 18 C; 15 G; 7 T; 0 other; |
| Query Match | 0.9%; Score 49.4; DB 22; Length 51; |
| Best Local Similarity | 98.0%; Pred. No. 0.0025; |
| Matches 50; Conservative | 0; Mismatches 1; Indels 0; Gaps 0 |
| OY | 200 CCCCTGGAAAGTCATCCCGCATCCGAGAGCCAGATCCCGCCCACTTG 250 |
| | |
| Db | 1 CCCCTGGAAAGTCATCCCGCATCCGAGAGCCAGATCCCGCCCACTTG 51 |
| RESULT 3 | |
| AAT28105 | |
| ID | AAT28105 standard; DNA: 39 BP. |
| AC | AAT28105: |
| XX | |
| DT | 31-DEC-1996 (first entry) |
| XX | |
| DE | Probe 10F1 isolated from fibroblasts. |
| XX | |
| KM | Polymerase chain reaction; PCR; primer; amplify; human; fibroblast; AIDS; |
| KM | enhanced differential display; EDD; mRNA preparation; senescent cell; |
| KM | quiescent cell; dividing cell; senescence-related gene; gene expression; |
| KM | non-senescent cell; age-related lipofuscin; retina; therapy; liver spot; |
| KM | donor tissue; senescent melanocyte; melanin; hypopigmentation; ss. |
| XX | |
| OS | Synthetic. |
| XX | |
| PN | WO9613610-A2. |
| XX | |
| PD | 09-MAY-1996. |
| XX | |
| PF | 24-AUG-1995; 95MO-US11230. |
| XX | |
| PR | 31-OCT-1994; 94US-0332420. |
| XX | |
| PA | (GERO-) GERON CORP. |
| XX | |
| PI | Feng J, Funk W, Hirsch KS, Linskens MHK, Villeponteau B; |
| PI | West MD; |
| XX | |
| DR | WPI; 1996-251464/25. |
| XX | |
| PT | Identifying, isolating and regulating senescence-related genes - |

PT useful to ameliorate problems associated with accumulation of
 PT senescent cells, e.g. age-related lipofuscin accumulation in the
 PT retina and AIDS
 XX
 PS Claim 8: Page 38; 135pp; English.
 XX
 CC AAT28076-T28113, and AAT28131-T28173 represent novel senescent-related
 CC gene sequences isolated from fibroblasts using the method of the
 CC invention. In the method of the invention, mRNA is isolated from a
 CC senescent cell, and a young quiescent cell, and the mRNAs are amplified
 CC (using primers such as those shown in AAT28044-T28075) in separate
 CC reaction mixtures. The amplified sequences are then separated by size or
 CC charge, and the products are analysed to identify a gene from young
 CC quiescent cells and dividing cells, that is present at a different level
 CC from senescent cells. To enhance the method even more, it can be
 CC performed in conjunction with an enhanced differential display (EDD)
 CC method (an mRNA preparation method) on the fibroblasts. The method can
 CC be used for the rapid and efficient identification and isolation of
 CC senescence-related genes and gene products, and to detect and distinguish
 CC between senescent and non-senescent cells. It can also be used to
 CC destroy cells expressing senescence specific (or related) gene products,
 CC and to screen for compounds capable of altering gene expression in
 CC senescent cells. The method can also be used to ameliorate problems
 CC associated with the accumulation of senescent cells such as age-related
 CC lipofuscin accumulation in the retina, and in the treatment of AIDS.
 CC Also, the method can be used to distinguish young cells from senescent
 CC cells in donor tissue, which is useful in removing senescent melanocytes
 CC overexpressing melanin which cause hypopigmentation, or liver spots.
 CC
 SQ Sequence 39 BP; 11 A; 7 C; 6 G; 15 T; 0 other;
 XX
 Query Match 0.7%; Score 39; DB 17; Length 39;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1412 AGATGATGATGTTACCCATTCACGACGATCTCTTT 1450
 DB 1 AGATGATGATGTTACCCATTCACGACGATCTCTTT 39
 XX
 RESULT 4
 AAAB7350/C
 ID AAAB7350 standard; DNA; 67 BP.
 AC
 XX AAAB7350;
 XX
 DT 08-JAN-2001 (first entry)
 XX
 DE Rat hepatocyte carcinogenesis biomarker nucleic acid SEQ ID NO:274.
 XX
 KW Rat; phenobarbital; carcinogenesis marker; carcinogenesis; detection;
 KW identification; carcinogenic; probe; primer; ds.
 XX
 OS Rattus norvegicus.
 XX
 PN WO200044902-A2.
 XX
 PD 03-AUG-2000.
 XX
 PF 28-JAN-2000; 2000WO-US00503.
 XX
 PR 29-JAN-1999; 99US-0118078.
 XX
 PA (SEAR) SEARLE & CO G D.
 XX
 PI Bunch RT, Curtis SW, Rodi CP, Morris DL;
 XX
 DR WPI: 2000-505977/45.
 XX
 PT New nucleic acid encoding a carcinogenic biomarker, induced by
 PT phenobarbital treatment of rat hepatocytes, useful for identifying
 PT carcinogenic compounds -
 XX

PS Claim 1: Page 137; 240pp; English.
 XX
 CC AAA87080 to AAA87656 represent nucleic acid sequences (NI) encoding a
 CC carcinogenesis biomarkers. The carcinogenesis biomarkers are induced by
 CC treating rat hepatocytes with phenobarbital. The nucleic acids are
 CC useful for identifying carcinogenic compounds. The nucleic acids molecules
 CC can be used to derive probes and/or primers for detecting or inducing
 CC carcinogenesis, respectively.
 CC
 SQ Sequence 67 BP; 25 A; 11 C; 10 G; 21 T; 0 other;
 XX
 Query Match 0.7%; Score 36.6; DB 21; Length 67;
 Best Local Similarity 83.1%; Pred. No. 8.9;
 Matches 54; Conservative 0; Mismatches 9; Indels 2; Gaps 1;
 OY 5136 TAATTCACACTCT--GAAGCTAATTGTTACTGATGATTTGTTTCATATAA 5193
 DB 67 TAATTCACACTCTGAGACCTAATTCGATTTGATTTGATTTGTTCTAATMAA 8
 XX
 OY 5194 GTGAA 5198
 DB 7 GTGGA 3
 XX
 RESULT 5
 ABNS3617
 ID ABNS3617 standard; DNA; 65 BP.
 XX
 AC ABNS3617;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:26365.
 XX
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.
 XX
 OS Mus musculus.
 XX
 PN WO200210449-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 20-JUL-2001; 2001WO-IB01903.
 XX
 PR 28-JUL-2000; 2000US-221607P.
 XX
 PR 02-MAY-2001; 2001US-287724P.
 XX
 PA (COMP-) COMPUSEN INC.
 XX
 PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
 XX
 DR WPI: 2002-257383/30.
 XX
 PT New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of
 PT a genome, useful for detecting tissue-, pathology-, and
 PT developmental-specific genes -
 XX
 PS Example 1: SEQ ID 26365; 47pp; English.
 XX
 CC The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises
 CC several oligonucleotides, each capable of hybridizing selectively to a
 CC set of messenger RNAs transcribed from a given transcription unit of
 CC the genome, which encodes one or more messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterizing the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini

from human lung comprising single exon nucleic acid probes having one of 12614 nucleic acid sequences mentioned in the specification, or their complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridize at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Kartagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a single exon probe open reading frame of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 83 BP; 26 A; 21 C; 20 G; 16 T; 0 other;

Query Match 0.6%; Score 33.8; DB 24; Length 83;
Best Local Similarity 64.9%; Pred. No. 59; Mismatches 27; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 4697 GCTGTGACCCCTGGGCAAGTCATTAAGTGGCTCAGTTTCTTCTGTAA 4756
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 81 GCTGTGATTTCTGAGCAAAATTCCTGGCCCTCTGAGGCTAAGCTTCCCTTGTAAG 22
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4757 ATGGGATATTAATTAAT 4773
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 21 ATGGGATTAATTAATTAAT 5

RESULT 8
AAT13984
ID AAT13984 standard; cDNA, 95 BP.
XX
AC AAT13984;
XX
DT 19-JUN-1996 (first entry)
XX
DE Stearoyl-CoA desaturase gene cDNA clone L36 consensus region.
XX
KW Obesity; body weight disorder; diagnosis; therapy;
KW gene modulator; fat; adipose; stearoyl-CoA desaturase; ss.
OS Mus musculus.
XX
PN W09605861-A1.
XX
PD 29-FEB-1996.
XX

PF 23-AUG-1995; 95MO-US10918.
XX
PR 06-JUN-1995; 95US-0470868.
PR 23-AUG-1994; 94US-0294522.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Tartaglia LA;
XX
DR WPI: 1996-151150/15.
XX
PT New gene which is differentially expressed in body weight disorders
PT -corresp. gene prod., antibodies and gene modulators; for
PT diagnosing and treating a weight disorder, such as obesity
XX
PS Example 6: Fig 3b; 162pp; English.
XX
CC 2 Regions (AAT13983 and AAT13984), complementary to bases 2-157 and
CC 123-217 of mouse cDNA L36 (AAT13982), show marked homology to
CC nucleotides 3505-3660 and 3447-3541, respectively, of the mouse
CC stearoyl-CoA desaturase gene SCD1, which encodes an enzyme that
CC converts saturated fats to mono-unsaturates. L36 cDNA was isolated
CC using a genetic obesity paradigm for the identification of genes
CC that are differentially expressed in obese vs. lean mice; levels
CC of L36 mRNA are higher in livers of obese mice than in livers of
CC their lean littermates. L36, and other identified genes (see
CC AAT13979-81 and AAT13985-T14008), can be used to diagnose body wt.
CC disorders.
XX

Sequence 95 BP; 24 A; 26 C; 14 G; 31 T; 0 other;

Query Match 0.6%; Score 31; DB 17; Length 95;
Best Local Similarity 57.9%; Pred. No. 37e+02;
Matches 55; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 4994 CAAATTTGACAAATACATTAACACACGCAATTCACGGAACATACAGATGCCCTTTC 5053
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 CAAGTGCACAACTACACCTACCCACTGATGAAACCATTCGAGAAACCTTCCCTTCC 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5054 TGTGATTGGTGGGATTTTTCCTTTTATGTGG 5088
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 CCTGATGACATTTTTCCTTTCCTTTTATGTGG 95
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
ABN67931/C
ID ABN67931 standard; DNA, 93 BP.
XX
AC ABN67931;
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polynucleotide SEQ ID NO 3775.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS Streptococcus agalactiae.
XX
PN W0200234771-A2.
XX
PD 02-MAY-2002.
XX
DE 29-OCT-2001; 2001WO-GB04789.
XX
PF 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
XX (GENO-) INST GENOMIC RES.
XX

| BASE COUNT | 24 a | 30 c | 23 g | 23 t |
|-----------------------|--|---|-------|--------------|
| Query Match | 1.5% | Score 79.4 | DB 12 | Length 100 |
| Best Local Similarity | 97.8% | Pred. No. 1.1e-05 | | |
| Matches | 91 | Conservative | 0 | Mismatches 1 |
| | | | | Indels 1 |
| | | | | Gaps 1 |
| Db | 9 | TTACACATTCTTCATGTTGATGGCGGCCCGCGCTGGCGCTATGACCGGACGAAGAAG | 1250 | |
| QY | 1251 | TTCTCCAAAGCGCCGATCTGTGGCCAGGATTAANA | 1283 | |
| Db | 68 | TTCTCCAAAGCGCCGATCTGTGGCCAGGATTAANA | 100 | |
| RESULT 3 | | | | |
| LOCUS | AI339479/c | | | |
| DEFINITION | q116d01.x1 NCI_CGAP_GC4 Homo sapiens CDNA clone IMAGE:1947745 3' | 94 bp | MRNA | linear |
| ACCESSION | AI339479 | | | |
| VERSION | AI339479.1 | | | |
| KEYWORDS | EST. | | | |
| SOURCE | human. | | | |
| ORGANISM | Homo sapiens | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| AUTHORS | Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo. | | | |
| TITLE | 1 (bases 1 to 94) | | | |
| JOURNAL | NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . | | | |
| COMMENT | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index | | | |
| | Unpublished (1997) | | | |
| | Contact: Robert Strausberg, Ph.D. | | | |
| | Email: gcraps-remail.nih.gov | | | |
| | Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael | | | |
| | Bumert-Buck, M.D., Ph.D. | | | |
| | CDNA Library Preparation: M. Bento Soares, Ph.D. | | | |
| | CDNA Library Arrayed by: Greg Lennon, Ph.D. | | | |
| | DNA Sequencing by: Washington University Genome Sequencing Center | | | |
| | clone distribution: NCI-CGAP clone distribution information can be | | | |
| | found through the I.M.A.G.E. Consortium/ILNL at: | | | |
| | www-bio.lnlnl.gov/db/rrp/image/image.html | | | |
| | Trace considered overall poor quality | | | |
| | Seq primer: -40UP from Glbco | | | |
| | High quality sequence stop: 1. | | | |
| FEATURES | Location/Qualifiers | | | |
| SOURCE | 1..94 | | | |
| | /organism="Homo sapiens" | | | |
| | /db_xref="taxon:9606" | | | |
| | /clone_image="1947745" | | | |
| | /clone_lib="NCI-CGAP_GC4" | | | |
| | /tissue_type="pooled germ cell tumors" | | | |
| | /lab_host="DH10B" | | | |
| | /note="Vector: p7T3D-Pac (Pharmacia) with a modified | | | |
| | polylinker; 1st strand cDNA was prepared from 3 pooled | | | |
| | germ cell tumors, and was then primed with a Not I - | | | |
| | oligo(dT) primer. Double-stranded cDNA was ligated to Eco | | | |
| | RuI adaptors (Pharmacia), digested with Not I and cloned | | | |
| | into the Not I and Eco RI sites of the modified p7T3 | | | |
| | vector. Library is normalized. Library was constructed by | | | |
| | Bento Soares and M. Fatima Bonaldo." | | | |
| BASE COUNT | 26 a | 25 c | 24 g | 19 t |
| ORIGIN | | | | |

| | | | |
|------------|---|---|------|
| OY | 992 | CGAATGCTGGGGGCTTAAGCCACCACCGGTGGGAGCAAGCGTGCCTCTTGGA | 1051 |
| Dd | 94 | CGAATGCTGGGGGCTTAAGCTCCTTCGCCCTGGGTGGTAACACTGTAGCGCCCTCATCTGA | 35 |
| OY | 1052 | TATCGCTTATGACAAAGACATTAGCCCCGGG | 1085 |
| Dd | 34 | TATCGCTTATGACAAAGACATTAGCCTCCGG | 1 |
| RESULT 4 | Bg113100 | | |
| LOCUS | Bg113100 | | |
| DEFINITION | 60228375FF1 NIH_MGC_86 Homo sapiens CDNA clone IMAGE:4371325 5' | | |
| ACCESSION | Bg113100 | | |
| VERSION | Bg113100.1 | GI:12606606 | |
| KEYWORDS | EST. | | |
| SOURCE | human. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. | | |
| JOURNAL | NIH-MGC http://mgc.nci.nih.gov/ | | |
| COMMENT | National Institutes of Health, Mammalian Gene Collection (MGC) | | |
| | Unpublished (1999) | | |
| | Contact: Robert Strausberg, Ph.D. | | |
| | Email: cgaabs-remail.nih.gov | | |
| | Tissue Procurement: ATCC | | |
| | CDNA Library Preparation: Life Technologies, Inc. | | |
| | CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) | | |
| | DNA Sequencing by: Incyte Genomics, Inc. | | |
| | Clone distribution: MGC clone distribution information can be | | |
| | found through the I.M.A.G.E. Consortium/LMNL at: | | |
| | http://image.llnl.gov | | |
| | plate: LMNL0029 row: P column: 14 | | |
| | High quality sequence stop: 73. | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..73 | | |
| | /organism="Homo sapiens" | | |
| | /db_xref="taxon:9606" | | |
| | /clone_image="4371325" | | |
| | /clone_id="NIH.MGC.86" | | |
| | /tissue_type="osteosarcoma, cell line" | | |
| | /lab_host="DH10B (phage-resistant)" | | |
| | /note="Organ: Bone; Vector: pCMV-SPORT6; Site_1: NotI; | | |
| | Site_2: SalI; Cloned unidirectionally; oligo-dT primed. | | |
| | Average insert size 1.533 kb. Library enriched for | | |
| | full-length clones and constructed by Life Technologies. | | |
| | Note: This is a NIH-MGC Library." | | |
| BASE COUNT | 12 a 20 c 19 g 22 t | | |
| ORIGIN | | | |
| | Query Match | 1.4%; Score 73; DB 12; Length 73; | |
| | Best Local Similarity | 100.0%; Pred. No. 0.00026; | |
| | Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | |
| OY | 3263 | GAGCTTATTTTCACGCCCCTCCTCGTGGTCAGAACGAGGGCAGTCGATGCCCCCT | 3322 |
| Dd | 1 | GGGTTATTTTCAGCCCTCTCTGTGGTCAGAACGAGGGCAGTCGATGCCCCCT | 60 |
| OY | 3323 | GCTTACTGTGTA 3335 | |
| Dd | 61 | GCTTACTGTGTA 73 | |
| RESULT 5 | AA501922 | 68 bp mRNA linear EST 18-AUG-1997 | |
| LOCUS | AA501922 | | |
| DEFINITION | nq954d08.s1 NCI_GCAP_L12 Homo sapiens CDNA clone IMAGE:938607, mRNA | | |
| ACCESSION | AA501922 | | |
| VERSION | AA501922.1 | GI:2236889 | |
| KEYWORDS | EST. | | |

| | |
|-----------------------|--|
| SOURCE | human. |
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. |
| AUTHORS | 1 (bases 1 to 68) |
| TITLE | NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . |
| JOURNAL | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index |
| COMMENT | Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: David E. Kleiner, M.D., Ph.D., Rodrigo F. Chuqui, M.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: David B. Kitzman, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: www-bio.livl.nih.gov/dbtrp/image/image.html Insert length: 607 Std Error: 0.00 Seq primer: -40m13 fwd. ET from Amersham High quality sequence stop: 51. Location/Qualifiers 1. 68 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:938607" /clone_11d="NCI_CGAP_L12" /sex="male" /tissue_type="liver" /lab_host="DH10B" /note="Vector: PAMP10; mRNA made from invasive hepatocellular carcinoma, CDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Kitzman et al. (1996) Cancer Research 56:5380-5383." |
| BASE COUNT | 12 a 20 c 17 g 19 t |
| ORIGIN | |
| Query Match | 1.3%; Score 66.4; DB 9; Length 68; |
| Best Local Similarity | 98.5%; Pred. No. 0.0059; |
| Matches | 67; Conservative 0; Mismatches 1; Indels 0; Gaps 0; |
| OY | 2344 GCAGCCTCCTTGTGTGATTCACAGGACGAGTGTGCTGCCAGGACGCTCCCTCC 3003 Db 1 GCAGCCCTCTGTGTGTATTCAGAGTGTGATGACACTTGTCTGCACGACGCTCCCTCC 60 |
| OY | 3004 TGCACACA 3011 |
| Db | 61 TGCACACA 68 |
| RESULT 6 | |
| AUI03573 | 50 bp mRNA linear EST 30-AUG-2001 |
| LOCUS | AUI03573 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone |
| DEFINITION | HEP3102, mRNA sequence. |
| ACCESSION | AUI03573 |
| VERSION | AUI03573.1 GI:13553094 |
| KEYWORDS | EST. |
| SOURCE | human. |
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. |
| AUTHORS | 1 (bases 1 to 50) |
| TITLE | Suzuki,Y., Talita,H., Tsunoda,T., Mizushima-Sugano,T., Seese,J., Hatakeyama,Y., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S. Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites EMBO Rep. 2 (5), 388-393 (2001) 21270072 Contact: Yutaka Suzuki |
| JOURNAL | |
| MEDLINE | |
| COMMENT | |

Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano, S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP13102"
/note="Sugano Homo sapiens cDNA library"

BASE COUNT
8 a 18 c 12 g 12 t
dimethylfumarate treated U937 cells"

ORIGIN
Query Match 1.0%; Score 50; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 ACGGTCACCGCTGGCCAGCTTAAATTCGCCGCTCGGGACCT 78
Db 1 ACGGTCACCGCTGGCCAGCTTAAATTCGCCGCTCGGGACCT 50

RESULT 7
AUI03578
LOCUS
DEFINITION
AUI03578 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
ADSE01417, mRNA sequence.
AUI03578
VERSION
AUI03578.1 GI:13553099
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 50)
Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki
Y., Nakamura, Y., Suyama, A. and Sugano, S.
Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
EMBO Rep. 2 (5), 388-393 (2001)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano, S.
Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ADSE01417"
/note="Sugano Homo sapiens cDNA library"

BASE COUNT
10 a 19 c 17 g 4 t
dimethylfumarate treated U937 cells"

ORIGIN
Query Match 1.0%; Score 50; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 ACGGCGCTAGCGCGACACAGCTTACGTCGACAGCGCGCGCTCAG 134
Db 1 ACGGCGCTAGCGCGACACAGCTTACGTCGACAGCGCGCGCTCAG 50

RESULT 8
AUI03571
LOCUS
DEFINITION
AUI03571 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP07904, mRNA sequence.
AUI03571
VERSION
AUI03571.1 GI:13553092
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 50)
Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki
Y., Nakamura, Y., Suyama, A. and Sugano, S.
Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
EMBO Rep. 2 (5), 388-393 (2001)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo

FEATURES
source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP07904"
/note="Sugano Homo sapiens cDNA library"

BASE COUNT
9 a 19 c 17 g 5 t
dimethylfumarate treated U937 cells"

ORIGIN
Query Match 0.9%; Score 48.4; DB 9; Length 50;
Best Local Similarity 98.0%; Pred. No. 32;
Matches 49; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 83 GCACCGCGCTAGCGCGACACAGCTTACGTCGACAGCGCGCGCTC 132
Db 1 GCACCGCGCTAGCGCGACACAGCTTACGTCGACAGCGCGCGCTC 50

RESULT 9
AUI03576
LOCUS
DEFINITION
AUI03576 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP1416, mRNA sequence.
AUI03576
VERSION
AUI03576.1 GI:13553097
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 50)
Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki
Y., Nakamura, Y., Suyama, A. and Sugano, S.
Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
EMBO Rep. 2 (5), 388-393 (2001)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo

FEATURES
source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP1416"
/note="Sugano Homo sapiens cDNA library"

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp
 Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
 S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source

1..50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HEP21416"
 /clone_lib="Sugano Homo sapiens cDNA library"
 /note="Differential display comparison of untreated and
 dimethylfumarate treated U937 cells"

BASE COUNT

8 a 19 c 14 g 9 t

ORIGIN

Query Match 0.9%; Score 48.4; DB 9; Length 50;
 Best Local Similarity 98.0%; Pred. No. 32;
 Matches 49; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 51 AGCCTTAATTCGGCTCGGAGACCTCCACGACGCGGTACGCCG 100
 |||||||
 DB 1 AGCCTTAATTCGGCTCGGAGACCTCCACGACGCGGTACGCCG 50

RESULT 10
 AUI03572 50 bp mRNA linear EST 30-AUG-2001
 LOCUS AUI03572 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone

DEFINITION HEP12765, mRNA sequence.
 ACCESSION AUI03572
 VERSION AUI03572.1 GI:13553093

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 50)
 Suzuki, Y., Talra, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
 H., Oka, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki
 Y., Nakamura, Y., Suyama, A. and Sugano, S.
 Diverse transcriptional initiation revealed by fine, large-scale
 mapping of mRNA start sites
 EMBO Rep. 2 (5), 388-393 (2001)
 21270072
 Contact: Yutaka Suzuki
 Department of Medical Science, University of Tokyo
 Institute of Medical Science, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp
 Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
 S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source

1..50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HEP12765"
 /clone_lib="Sugano Homo sapiens cDNA library"
 /note="Differential display comparison of untreated and
 dimethylfumarate treated U937 cells"

BASE COUNT

8 a 16 c 12 g 14 t

ORIGIN

Query Match 0.9%; Score 46.8; DB 9; Length 50;
 Best Local Similarity 96.0%; Pred. No. 67;
 Matches 48; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 29 AGGTCACCGCTGCGACCTTAAGCTTAATCCCGCGCTGGGACCT 78
 |||||||
 DB 1 AGGTCACCGCTGCGACCTTAAGCTTAATCCCGCGCTGGGACCT 50

RESULT 11

AA581009

LOCUS

DEFINITION

AA581009 58 bp mRNA linear EST 05-JAN-1998
 nc83106.r1 NCI-CGAP_GCI Homo sapiens cDNA clone IMAGE:797411 5'
 similar to SW:AC02_MOUSE P13011 ACYL-COA DESATURASE 2 ; mRNA
 sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 58)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies Inc., David Krizman,
 Ph.D.
 cDNA Library Arraying: Greg Lennon, Ph.D.
 DNA sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.lnl.gov/brf/image/image.html
 Seq primer: -28ml3 rev1 ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers

FEATURES

source

1..58
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:797411"
 /clone_lib="NCI-CGAP_GCI"
 /tissue_type="bulk germ cell seminoma"
 /lab_host="DH10B"
 /note="Vector: PCWV-SPORT2; Site 1: SalI; Site 2: NotI;
 cloned unidirectionally. Primer: Oligo dr."

BASE COUNT

9 a 14 c 21 g 14 t

ORIGIN

Query Match 0.9%; Score 45.8; DB 9; Length 58;
 Best Local Similarity 87.7%; Pred. No. 1e+02;
 Matches 50; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1003 GGTGCTTAAGCCACCTGGCTGGTGAACAGTCTGCCACCTCTTGGGATATGCTCC 1059
 |||||||
 DB 1 GGTGCTTAAGCCACCTGGTGGTGAAGAGTCTGCCACCTCTTGGGATATGCTCC 57

RESULT 12

AA108264

LOCUS

DEFINITION

AA108264 83 bp mRNA linear EST 05-SEP-1997
 EST0007 rat lambda ZAPIT library (C.P.Hamel) Rattus norvegicus cDNA
 clone PC038 5' similar to Stearyl-CoA denaturase, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

1 (bases 1 to 83)
 Soto-Prior, A., Lavigne-Rebillard, M., Lenoir, M., Ripoll, C.,
 Rebillard, G., Vago, P., Pujol, R. and Hamel, C.P.
 Identification of preferentially expressed cochlear genes by
 systematic sequencing of a rat cochlea cDNA library
 Brain Res. Mol. Brain Res. 47 (1-2), 1-10 (1997)
 97364947

| | |
|-----------------------|---|
| KEYWORDS | EST. |
| SOURCE | house mouse. |
| ORGANISM | Mus musculus |
| REFERENCE | Mammalia; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi; Eumariola; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |
| AUTHORS | 1 (bases 1 to 80) Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R. |
| TITLE | The WashU-HHMI Mouse EST Project |
| JOURNAL | Unpublished (1996) |
| COMMENT | Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watsn.wustl.edu This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MG1:494899 putative full length read vector to vector length is 101. |
| FEATURES | Location/Qualifiers 1..80 /organism="Mus musculus" /strain="C57BL/6J x DBA/2J F1" /db_xref="taxon:10090" /clone="IMAGE:834683" /clone_lib="Knowles Solter mouse blastocyst B3" /tissue_type="Blastocyst" /dev_stage="embryo (pre-implantation)" /lab_host="DH10B" /note="Organ: embryo; Vector: pSPORT; Site:1: NotI; Site 2: SalI; Cloned unidirectionally from mRNA prepared from 800 blastocysts. Primer: SalI(dT): 5'-GGGTGACACCTGACACCTTTTCTTTTCTTT-3'. cDNAs were cloned into the NotI/SalI sites of a pSPORT vector (Life Technologies). Two different size selections: B1 (larger inserts) and B3." |
| BASE COUNT | 24 a 12 c 16 g 28 t |
| ORIGIN | |
| Query Match | 0.79; Score 38.2; DB 9; Length 80; |
| Best Local Similarity | 74.7%; Pred. No.3.4e+03; |
| Matches | 62; Conservative 0; Mismatches 18; Indels 3; Gaps 1; |
| OY | 1384 CTACTGATTAATGCTACACAGATGCTAAAGATGATGTTAACCATTCAGTACAGTA 1443 |
| DB | 1 CAATGATATTGCTATCGGGGTGTTA--ANGATGATTTAACTATTCCGGTACAGTA 57 |
| OY | 1444 TTCTTTTAAATTCAAAAGTATT 1466 |
| DB | 58 TTCTTTTAAATTCAGAAAGCTTT 80 |

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 20:58:57 ; Search time 116 Seconds

(Without alignments)
13803.098 Million cell updates/sec

Title: US-09-918-187-3

Perfect score: 5221

Sequence: 1 ataaagggggcgtgggaaa.....aatctaaaaaaaaaaaaa 5221

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 684418

Minimum DB seq length: 8

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCITUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 39 | 0.7 | 39 | 1 | US-08-332-420-62 |
| 2 | 28.8 | 0.6 | 50 | 1 | US-08-222-177A-367 |
| 3 | 28.8 | 0.6 | 65 | 1 | US-08-222-177A-116 |
| 4 | 28.8 | 0.6 | 65 | 1 | US-08-222-177A-421 |
| 5 | 28.2 | 0.5 | 51 | 1 | US-08-222-177A-337 |
| 6 | 27.8 | 0.5 | 57 | 1 | US-08-222-177A-382 |
| 7 | 27.4 | 0.5 | 75 | 3 | US-09-284-782-20 |
| 8 | 27.2 | 0.5 | 50 | 1 | US-08-222-177A-328 |
| 9 | 27 | 0.5 | 100 | 1 | US-08-145-705A-17 |
| 10 | 26.8 | 0.5 | 46 | 1 | US-08-222-177A-71 |
| 11 | 26.8 | 0.5 | 49 | 1 | US-08-222-177A-361 |
| 12 | 26.8 | 0.5 | 53 | 1 | US-08-222-177A-304 |
| 13 | 26.8 | 0.5 | 56 | 1 | US-08-222-177A-412 |
| 14 | 26.8 | 0.5 | 65 | 1 | US-08-222-177A-134 |
| 15 | 26.8 | 0.5 | 91 | 1 | US-08-222-177A-107 |
| 16 | 26.6 | 0.5 | 89 | 1 | US-08-222-177A-98 |
| 17 | 26.4 | 0.5 | 56 | 1 | US-08-222-177A-394 |
| 18 | 26.4 | 0.5 | 62 | 6 | 5506118-3 |
| 19 | 26.4 | 0.5 | 69 | 4 | US-09-269-911A-8 |
| 20 | 26.4 | 0.5 | 84 | 3 | US-09-284-782-26 |
| 21 | 26.2 | 0.5 | 51 | 1 | US-08-222-177A-89 |
| 22 | 26.2 | 0.5 | 51 | 1 | US-08-222-177A-317 |
| 23 | 26 | 0.5 | 83 | 1 | US-08-120-827-100 |
| 24 | 26 | 0.5 | 83 | 1 | US-08-478-675-100 |
| 25 | 26 | 0.5 | 92 | 5 | PCT-US93-06251-39 |
| 26 | 25.8 | 0.5 | 72 | 1 | US-08-222-177A-131 |
| 27 | 25.8 | 0.5 | 72 | 1 | US-08-222-177A-427 |

| | | | | | | |
|----|------|-----|----|---|--------------------|-------------------|
| 28 | 25.8 | 0.5 | 95 | 1 | US-08-518-878B-3 | Sequence 3, Appl1 |
| 29 | 25.8 | 0.5 | 95 | 1 | US-08-294-522B-3 | Sequence 3, Appl1 |
| 30 | 25.8 | 0.5 | 95 | 2 | US-08-807-861A-3 | Sequence 3, Appl1 |
| 31 | 25.8 | 0.5 | 95 | 2 | US-08-470-868A-3 | Sequence 3, Appl1 |
| 32 | 25.8 | 0.5 | 95 | 3 | US-09-210-681-3 | Sequence 3, Appl1 |
| 33 | 25.8 | 0.5 | 95 | 3 | US-08-946-719A-3 | Sequence 3, Appl1 |
| 34 | 25.6 | 0.5 | 51 | 1 | US-08-222-177A-391 | Sequence 391, App |
| 35 | 25.6 | 0.5 | 62 | 1 | US-08-222-177A-56 | Sequence 56, App |
| 36 | 25.6 | 0.5 | 80 | 4 | US-09-284-627-15 | Sequence 15, App |
| 37 | 25.4 | 0.5 | 40 | 1 | US-08-222-177A-157 | Sequence 157, App |
| 38 | 25.4 | 0.5 | 45 | 1 | US-08-222-177A-157 | Sequence 157, App |
| 39 | 25.4 | 0.5 | 54 | 1 | US-08-469-802B-27 | Sequence 27, App |
| 40 | 25.4 | 0.5 | 54 | 2 | US-08-267-803B-45 | Sequence 45, Appl |
| 41 | 25.4 | 0.5 | 75 | 2 | US-08-776-944-13 | Sequence 13, Appl |
| 42 | 25.2 | 0.5 | 38 | 1 | US-08-222-177A-198 | Sequence 198, App |
| 43 | 25.2 | 0.5 | 39 | 1 | US-08-222-177A-137 | Sequence 137, App |
| 44 | 25.2 | 0.5 | 40 | 1 | US-08-222-177A-119 | Sequence 119, App |
| 45 | 25.2 | 0.5 | 40 | 1 | US-08-222-177A-400 | Sequence 400, App |

ALIGNMENTS

RESULT 1
US-08-332-420-62
Sequence 62, Application US/08332420
Patent No. 5744300
GENERAL INFORMATION:
APPLICANT: Maarten H.K. Linckens, et al.
TITLE OF INVENTION: METHODS AND REAGENTS FOR THE
TITLE OF INVENTION: IDENTIFICATION AND REGULATION
NUMBER OF INVENTION: OF SENSENCE-RELATED GENES
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,420
FILING DATE: October 31, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/235,180
FILING DATE: April 29, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 210/041
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 39
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-332-420-62
Query Match 0.7%; Score 39; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1412 AGATGATGATGTTTACCCATTCCAGTACGATTCCTTTT 1450


```
STATE: wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 304:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd94rs
US-08-222-177A-304

Query Match          0.5%; Score 26.8; DB 1; Length 53;
Best Local Similarity 81.6%; Pred. No. 3.1e+02;
Matches 31; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Cy 4078 TGTATGTCGCTGGCTGCTGGCTGGAGTGTCTG 4115
      ||| ||||| ||| ||||| ||| ||||| |||
Db 45 TGTGTCGTCTGCTGTCTGTCTGTCTGTCTGTCTG 8

RESULT 13
US-08-222-177A-412/C
Sequence 412, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Rose & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
```

```
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 412:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-222-177A-412

Query Match          0.5%; Score 26.8; DB 1; Length 56;
Best Local Similarity 81.6%; Pred. No. 3.2e+02;
Matches 31; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Cy 4078 TGTATGTCGCTGGCTGCTGGCTGGAGTGTCTG 4115
      ||| ||||| |||| ||||| ||| ||||| |||
Db 38 TGTATGTCGTCTGTCTGTCTGTCTGTCTGTCTG 1

RESULT 14
US-08-222-177A-134/C
Sequence 134, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Rose & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd28rs
US-08-222-177A-134

Query Match          0.5%; Score 26.8; DB 1; Length 65;
Best Local Similarity 81.6%; Pred. No. 3.5e+02;
```


